

#4
520.38682X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

JC866 U.S. PTO
09/608151
06/30/00


Applicant(s): Shinichi NAKAYAMA, ET AL.

Serial No.:

Filed: June 30, 2000

Title: STORAGE SUBSYSTEM AND STORAGE CONTROLLER

Group:

LETTER CLAIMING RIGHT OF PRIORITY

Honorable Commissioner of
Patents and Trademarks
Washington, D.C. 20231

June 30, 2000

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Patent Application No.(s) 11-353806 filed December 14, 1999.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP


Alan E. Schiavelli
Registration No. 32,087

AES/rdh
Attachment
(703) 312-6600

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application:

1999年12月14日

出願番号
Application Number:

平成11年特許願第353806号

出願人
Applicant(s):

株式会社日立製作所

JCA0609/608151 PRO
06/30/00

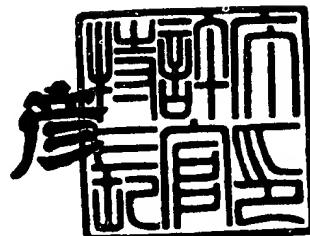


CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年5月19日

特許長官
Commissioner,
Patent Office

近藤 隆



【書類名】 特許願
【整理番号】 PNT991022
【提出日】 平成11年12月14日
【あて先】 特許庁長官 殿
【国際特許分類】 G06F 3/06
【発明者】
【住所又は居所】 神奈川県小田原市国府津2880番地 株式会社日立製作所
【氏名】 中山 信一
【発明者】
【住所又は居所】 神奈川県小田原市国府津2880番地 株式会社日立製作所
【氏名】 横畠 静生
【特許出願人】
【識別番号】 000005108
【氏名又は名称】 株式会社日立製作所
【代理人】
【識別番号】 100068504
【弁理士】
【氏名又は名称】 小川 勝男
【電話番号】 03-3661-0071
【選任した代理人】
【識別番号】 100086656
【弁理士】
【氏名又は名称】 田中 恭助
【電話番号】 03-3661-0071
【手数料の表示】
【予納台帳番号】 081423
【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1
【物件名】 図面 1
【物件名】 要約書 1
【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 記憶サブシステム及び記憶制御装置

【特許請求の範囲】

【請求項1】

上位外部から受け取った入出力要求に応じて下位外部との間で入出力データの転送を制御する記憶制御装置において、前記記憶制御装置は、上位外部とのインターフェースに応じて前記入出力要求を受け取る少なくとも1台の外部インターフェースコントローラと、前記入出力要求の処理をする少なくとも1台の制御プロセッサと、前記外部インターフェースコントローラと前記制御プロセッサとの間に介在し、両者間の情報伝送の伝送路となるファイバチャネルインターフェースのループとを有することを特徴とする記憶制御装置。

【請求項2】

前記外部インターフェースコントローラの上位外部とのインターフェースは、ファイバチャネルインターフェースであることを特徴とする請求項1記載の記憶制御装置。

【請求項3】

前記外部インターフェースコントローラは、ファイバチャネルインターフェースと異なる上位外部とのインターフェースと前記ループのファイバチャネルインターフェースとの間のインターフェース変換機能を有することを特徴とする請求項1記載の記憶制御装置。

【請求項4】

前記ループは、入力信号に応じて前記外部インターフェースコントローラと前記制御プロセッサとの間の伝送経路を形成する電子スイッチ機構を有することを特徴とする請求項1記載の記憶制御装置。

【請求項5】

上位外部から受け取った入出力要求に応じて下位側の記憶媒体の駆動装置との間で入出力データの転送を制御する記憶サブシステムにおいて、前記記憶サブシステムは、上位外部とのインターフェースに応じて前記入出力要求を受け取る少なくとも1台の外部インターフェースコントローラと、前記入出力要求の処理をする

少なくとも1台の制御プロセッサと、前記外部インターフェースコントローラと前記制御プロセッサとの間に介在し、両者間の情報伝送の伝送路となるファイバチャネルインターフェースのループとを有することを特徴とする記憶サブシステム。

【請求項6】

上位外部から受け取った入出力要求に応じて下位外部との間で入出力データの転送を制御する記憶制御装置において、前記記憶制御装置は、前記入出力要求を受け取る複数の外部インターフェースコントローラと、前記入出力要求の処理をする複数台の制御プロセッサと、前記外部インターフェースコントローラと前記制御プロセッサとの間に介在し、両者間の情報伝送の伝送路となるファイバチャネルインターフェースのループとを有し、前記制御プロセッサは、前記ループを介して送られる前記入出力要求のうち自プロセッサ宛てのアドレスを有するフレームを読み込む手段と、読み込んだフレームについて前記入出力要求の処理をする手段とを有することを特徴とする記憶制御装置。

【請求項7】

上位外部から受け取った入出力要求に応じて下位外部との間で入出力データの転送を制御する記憶制御装置において、前記記憶制御装置は、前記入出力要求を受け取る複数の外部インターフェースコントローラと、前記入出力要求の処理をする複数台の制御プロセッサと、前記外部インターフェースコントローラと前記制御プロセッサとの間に介在し、両者間の情報伝送の伝送路となるファイバチャネルインターフェースのループと、前記制御プロセッサによって共通にアクセスされ前記制御プロセッサ対応に処理すべき入出力要求の論理ユニット番号を記憶する記憶手段とを有し、前記制御プロセッサは、他の制御プロセッサの稼動状態を監視する手段と、他の制御プロセッサが停止していることを検出したとき、前記記憶手段上の前記論理ユニット番号を書き換えることにより、停止した制御プロセッサに関する処理を引き継ぐ手段とを有することを特徴とする記憶制御装置。

【請求項8】

前記記憶手段は、前記制御プロセッサ対応に前記ループ上の物理アドレスと処理すべき入出力要求の論理ユニット番号とを記憶し、前記制御プロセッサは、他の制御プロセッサが停止していることを検出したとき、前記記憶手段上の前記物

理アドレスと前記論理ユニット番号とを引き継ぐべく書き換えることにより、停止した制御プロセッサに関する処理を引き継ぐ手段とを有することを特徴とする請求項7記載の記憶制御装置。

【請求項9】

上位外部から受け取った入出力要求に応じて下位外部との間で入出力データの転送を制御する記憶制御装置において、前記記憶制御装置は、前記入出力要求を受け取る複数の外部インターフェースコントローラと、前記入出力要求の処理をする複数台の制御プロセッサと、前記外部インターフェースコントローラと前記制御プロセッサとの間に介在し、両者間の情報伝送の伝送路となるファイバチャネルインターフェースのループと、前記制御プロセッサによって共通にアクセスされ前記制御プロセッサ対応に処理すべき入出力要求の論理ユニット番号を記憶する記憶手段とを有し、前記制御プロセッサは、処理入出力要求数を計数する手段と、他の制御プロセッサに処理入出力要求数を通知する手段と、他の制御プロセッサの処理入出力要求数を取得する手段と、制御プロセッサ間の処理入出力要求数が平均化するように前記記憶手段上の前記論理ユニット番号を書き換える手段とを有することを特徴とする記憶制御装置。

【請求項10】

上位のホストコンピュータから受け取った入出力要求に応じて下位側の記憶媒体の駆動装置との間で入出力データの転送を制御する記憶サブシステムにおいて、前記記憶サブシステムは、上位ホストコンピュータとのインターフェースに応じて前記入出力要求を受け取る少なくとも1台の外部インターフェースコントローラと、データを一時的に格納するキャッシュメモリと、前記入出力要求を解析し、前記ホストコンピュータと前記キャッシュメモリとの間で行う入出力データの転送を制御する少なくとも1台の上位側の制御プロセッサと、前記外部インターフェースコントローラと前記上位側の制御プロセッサとの間に介在し、両者間の情報伝送の伝送路となるファイバチャネルインターフェースのループと、前記キャッシュメモリと前記記憶媒体の駆動装置との間で行う入出力データの転送を制御する少なくとも1台の下位側の制御プロセッサと、前記下位側の制御プロセッサと前記記憶媒体の駆動装置との間に介在し、下位駆動装置とのインターフェースに応じ

て前記記憶媒体の駆動装置との間で入出力データの転送を行うドライブインターフェースコントローラとを有することを特徴とする記憶サブシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、上位側でホストコンピュータと接続する記憶サブシステム及び記憶制御装置に係わり、特に処理性能及び信頼性の向上を図る記憶サブシステム及び記憶制御装置に関する。

【0002】

【従来の技術】

近年、コンピュータシステムの大規模化、データ処理の高速化、24時間あるいは365日無停止運転の必要性、データ転送インターフェースの高速化などに伴い、記憶制御装置に対して性能、信頼性、可用性の向上が強く要求されている。性能向上を目的とする記憶制御装置の例として、記憶制御装置内に内部ネットワークを備えることにより性能向上を図ることを目的とする特開平11-7359号公報に開示される技術がある。

【0003】

また記憶制御装置に接続するホストコンピュータとの間のインターフェースについて着眼すると、図8に示すように複数種類のインターフェースをもつホストコンピュータを接続する必要が生じている。記憶制御装置内のホストインターフェース部は、ホストインターフェースごとに設けられ、その制御プロセッサは、ホストコンピュータから受け取った入出力コマンドを解析し、信号線を介してキャッシュメモリ215との間でデータ転送を行う。例えば特開平9-325905号公報はこのような記憶制御装置を開示する。

【0004】

さらに近年、性能及び可用性を向上させるために、ホストコンピュータと記憶制御装置との間のインターフェースをSCSI(Small Computer System Interface)からファイバチャネルインターフェースに置き換えた公知技術が知られている。例えば特開平10-333839号公報は、記憶制御装置とホストコンピュータ

間をファイバチャネルインターフェースによって接続する技術を開示する。これはファイバチャネルインターフェースを有するホストコンピュータ専用の記憶制御装置である。

【0005】

【発明が解決しようとする課題】

上記の特開平11-7359号公報および特開平9-325905号公報に開示される技術は、ホストコンピュータから受け取った入出力要求を一つの制御プロセッサが処理するため、制御プロセッサの性能によって記憶制御装置全体の性能が押さえられるという問題がある。また制御プロセッサの障害によって関連するホストコンピュータから記憶制御装置が使用できなくなるという問題がある。特に昨今のファイバチャネルは100MB/Sという高速データ転送が可能であるため、制御プロセッサの処理性能がネックとなってファイバチャネルのもつデータ転送速度を充分生かせない。

【0006】

さらに特開平10-333839号公報に開示された技術はファイバチャネルインターフェース専用の記憶制御装置であるため、SCSIインターフェースを有するホストコンピュータを接続することができない。

【0007】

本発明の目的は、記憶制御装置の性能向上、特にファイバチャネルのもつ高速データ転送を生かすような高性能をもつとともに、信頼性、可用性の高い記憶サブシステム及び記憶制御装置を提供することにある。

【0008】

本発明の他の目的は、複数種類のインターフェースをもつホストコンピュータを接続可能な記憶サブシステム及び記憶制御装置を提供することにある。

【0009】

【課題を解決するための手段】

本発明は、上位外部から受け取った入出力要求に応じて下位側の記憶媒体の駆動装置との間で入出力データの転送を制御する記憶サブシステムにおいて、この記憶サブシステムは、上位外部とのインターフェースに応じて入出力要求を受け取

る少なくとも1台の外部インターフェースコントローラと、入出力要求の処理をする少なくとも1台の制御プロセッサと、外部インターフェースコントローラと制御プロセッサとの間に介在し、両者間の情報伝送の伝送路となるファイバチャネルインターフェースのループとを有する記憶サブシステムを特徴とする。またこのよきな記憶制御装置を特徴とする。

【0010】

また本発明は、上位外部とのインターフェースがファイバチャネルインターフェースであるような外部インターフェースコントローラ、およびファイバチャネルインターフェースと異なる上位外部とのインターフェースとファイバチャネルインターフェースとの間のインターフェース変換機能を有する外部インターフェースコントローラを設けた記憶サブシステム及び記憶制御装置を特徴とする。

【0011】

【発明の実施の形態】

以下、本発明の実施形態について図面により詳細に説明する。

【0012】

図1は、本発明の一実施形態を示すディスクサブシステムを含むシステムの構成図である。ディスク制御装置107は、上位外部側で、ホストコンピュータ100, 101及び102と接続する。ホストコンピュータ101は、メインフレーム系のコンピュータであり、メインフレーム系チャネルを介してディスク制御装置107と接続する。ホストコンピュータ100は、オープンシステム系のホストコンピュータであり、ファイバチャネルインターフェースを介してディスク制御装置107と接続する。ホストコンピュータ102は、オープンシステム系のコンピュータであり、SCSI (Small Computer System Interface)を介してディスク制御装置107と接続する。ディスク制御装置107は、下位外部側でファイバチャネルインターフェースのループ125及び126を介してドライブ127, 128, 129及び130と接続する。

【0013】

ホストインターフェースコントローラ(HIFC)103, 104及び105は、それぞれホストコンピュータ100, 101及び102と接続し、また他方で

ファイバチャネルインタフェースに準拠するループ133と接続する。制御プロセッサ114, 115, 116及び117は、一方でループ133と接続し、また他方で共通バス118と接続する。共通バス118には、制御プロセッサ114～117のほかに共通制御メモリ112、キャッシュメモリ122、制御プロセッサ119及び120が接続される。制御プロセッサ119および120は、それぞれファイバチャネル141を介してドライブインターフェースコントローラ(DIFC)123及び124と接続する。DIFC123および124は、それぞれループ125及び126を介してドライブ127, 128, 129及び130と接続する。制御プロセッサ114, 115, 116及び117は、信号線132を介してサービスプロセッサ131と接続する。

【0014】

HIFC103は、上位外部との間のインターフェースコントローラであり、ホストコンピュータ100からフレームの形式で受け取った入出力コマンド、データ及び制御情報をそのままの形式でループ133を介して制御プロセッサ114～117のいずれかへ転送する。またループ133を介して制御プロセッサ114～117からフレームの形式で受け取ったデータ及び制御情報をそのままホストコンピュータ100へ転送する。HIFC104は、ホストコンピュータ101から受け取ったチャネルコマンド、データ及び制御情報をファイバチャネルのフレーム形式に変換し、ループ133を介して制御プロセッサ114～117のいずれかへ転送する。また制御プロセッサ114～117からフレームの形式で受け取ったデータ及び制御情報をメインフレーム系のチャネルインタフェースに従うデータ形式に変換してホストコンピュータ101へ転送する。HIFC105は、ホストコンピュータ102から受け取った入出力コマンド、データ及び制御情報をファイバチャネルのフレーム形式に変換し、ループ133を介して制御プロセッサ114～117のいずれかへ転送する。また制御プロセッサ114～117からフレームの形式で受け取ったデータ及び制御情報をSCSIに従うデータ形式に変換してホストコンピュータ102へ転送する。1台のHIFC103, 104又は105にそれぞれ複数台のホストコンピュータ100, 101及び102を接続することも可能である。

【0015】

キャッシュメモリ122は、共通バス118のバスインターフェースを介してすべての制御プロセッサ114～117, 119及び120からアクセス可能なメモリであり、ホストコンピュータ100～102から送られたデータ及びドライブ127～130から読み出したデータを一時的に格納するために用いられる。キャッシュメモリ122上のデータは、キャッシュスロットと呼ばれるデータ管理単位に分割されている。

【0016】

共通制御メモリ112は、共通バス118を介してすべての制御プロセッサ114～117, 119及び120からアクセス可能な共通メモリであり、制御プロセッサ間の通信のための領域、キャッシュスロット管理テーブルなどのほかに、制御プロセッサ114～117の各々がループ133を介して取り込むべきフレームを設定するF C A L (Fibre Channel Arbitrated Loop)管理情報113を格納する。

【0017】

制御プロセッサ114～117の各々は、共通制御メモリ112上のF C A L 管理情報113を参照し、ループ133上を流れるフレームのうち設定されたアドレスをもつフレームを取り込み、入出力コマンドによって指定された入出力要求を実行する。すなわちリードコマンドの場合には、キャッシュメモリ122上に要求されたデータがあれば読み出して、ループ133及びH I F C 103～105のいずれかを介して要求元のホストコンピュータへ送信する。キャッシュメモリ122上に要求されたデータがなければ、制御プロセッサ119及び120に対して入出力要求を通知する。またライトコマンドの場合には、書き込みデータをキャッシュメモリ122上のキャッシュスロットに書き込み、制御プロセッサ119及び120に対して入出力要求を通知する。

【0018】

制御プロセッサ119及び120は、制御プロセッサ114～117から入出力要求の通知を受け取り、リードコマンドの場合にはドライブ127～130から要求されたデータを読み出してキャッシュメモリ122上のキャッシュスロッ

トに書き込む。またライトコマンドの場合には、キャッシュメモリ122上のデータをドライブ127～130上に書き込む。

【0019】

図2は、HIFC103～106と制御プロセッサ114～117との間に介在し、両者間の伝送路となるループ133及び関連機構の構成を示す図である。ループ133は、PBC(Port Bypass Circuit)108, 109, 110及び111を有し、いわゆるハブ(HUB)構造を形成している。PBC108～111は、各々1入力n出力の電子スイッチであり、図示するようにHIFC103～106及び制御プロセッサ114～117と接続するとともに、PBC108、PBC111間及びPBC109、PBC110間を接続する。本例のPBC108～111は、1入力2出力のスイッチであり、PBCに入力信号を与えると出力経路を限定することができる。ファイバコントローラ(FC)151は、制御プロセッサ114～117の各々に前置するファイバチャネルインタフェースコントローラであり、ループ133を介して送られるフレームの宛先アドレスを認識し、あらかじめ設定されたアドレスを宛先アドレスとするフレームを取り込んで接続される制御プロセッサへ送る。また制御プロセッサ114～117から受け取ったデータ及び制御情報をフレーム形式のデータにしてループ133へ送出する。ループ133は、HIFC103～106、FC151及び制御プロセッサ114～117を端末としてFibre Channel Arbitrated Loop(FCAL)と呼ばれるトポジカルなループ伝送路を形成する。ファイバチャネルの通信プロトコルについては、例えばANSI公開マニュアル「FIBRE CHANNEL PHYSICAL AND SIGNALLING (FC-PH)REV.4.3」に記載されている。

【0020】

例えばPBC108は、HIFC103を介してホストコンピュータ100と接続され、制御プロセッサ114, 115及びPBC111と接続可能である。従ってホストコンピュータ100から送られた入出力要求のコマンドは、PBC108を介して制御プロセッサ114, 115又はPBC111を介して制御プロセッサ116, 117で処理することができる。同様にホストコンピュータ101から送られた入出力要求のコマンドは、PBC109を介して制御プロセッ

サ114, 115又はPBC110を介して制御プロセッサ116, 117で処理することができる。

【0021】

本実施形態では、ファイバチャネル141, ループ125及び126にも、ファイバチャネルインタフェースを採用している。従って図示していないが、制御プロセッサ119, 120とファイバチャネル141との間にFC151が介在する。

【0022】

図3は、F CAL管理情報113のデータ構成を示す図である。F CAL管理情報113は、制御プロセッサ114～117の各々についてループ133を介して取り込むべきフレームと入出力処理の対象とするデバイス番号の範囲を設定するテーブルである。F CAL管理情報113の各エントリは、制御プロセッサ201、AL-PA(Arbitrated Loop Physical Address)202及びLUN(Logical Unit Number)203を有する。制御プロセッサ201は、制御プロセッサ114～117のいずれかの制御プロセッサの識別子である。AL-PA202は、その制御プロセッサに割り当てられたループ133上のアドレスである。LUN203は、制御プロセッサが入出力処理を分担する論理デバイス番号又は論理デバイス番号の範囲である。F CAL管理情報113の情報は、サービスプロセッサ131からの指示により設定又は設定解除することができる。

【0023】

図4は、制御プロセッサ114～117の処理の流れを示すフローチャートである。制御プロセッサ114～117は、各々周期的にF CAL管理情報113から自プロセッサのエントリを読み込み、接続されているFC151に自プロセッサのAL-PAを設定し、変更があればAL-PAを再設定する。FC151は、HIFC103及びループ133を介してホストコンピュータ100から送られるフレーム上のAL-PAを読み取り(ステップ301)、接続されている制御プロセッサのAL-PAでなければ(ステップ302NO)、処理を終了する。接続されている制御プロセッサのAL-PAであれば(ステップ302YES)、その制御プロセッサに通知する。通知を受けた制御プロセッサ114～117の

いずれかは、FC151を介してフレームを読み込み(ステップ303)、フレーム中の入出力コマンドのLUNがLUN203の範囲にあるか否かをチェックする(ステップ304)。もし指定されたLUNがLUN203の範囲になければホストコンピュータ100へエラー応答を返す。次に制御プロセッサは受け取った入出力コマンドに従って入出力要求を実行する(ステップ305)。

【0024】

制御プロセッサ114～117は、入出力要求が書き込み要求であれば、ホストコンピュータ100からデータを受領し、キャッシュメモリ122上の該当するキャッシュスロットにデータを書き込み、書き込み要求処理を終了する。データを書き込むべきキャッシュスロットのスロット番号はデータに付されたLBA(Logical Block Address)から計算される。またスロット番号に対応するキャッシュメモリ122上のメモアドレスは、共通制御メモリ112上のキャッシュスロット管理テーブルから得られる。入出力要求が読み出し要求であれば、キャッシュメモリ122上に要求されたデータがある場合、キャッシュメモリ122から読み出し、ループ133及びHIFC103を介してホストコンピュータ100へ送る。キャッシュスロット管理テーブルを参照することによって目的のデータの有無を判定することができる。キャッシュメモリ122上に要求されたデータがなければ、共通制御メモリ112上のプロセッサ間連絡領域に読み出し要求を書き込み、目的のデータがキャッシュメモリ122上に格納されたと判定したとき、キャッシュメモリ122から読み出してホストコンピュータ100へ送る。

【0025】

制御プロセッサ119, 120は、キャッシュメモリ122を探索し、キャッシュスロットにドライブ127～130に書き込むべきデータがあれば、ファイルチャネル141、DIFC123, 124及びループ125, 126を介してそのデータをドライブ127～130に書き込む。この書き込みは、ホストコンピュータ100と制御プロセッサ114～117との間の入出力要求の処理動作とは非同期に行われる。制御プロセッサ119, 120は、指定されたLUN及びLBAを物理デバイス番号と物理アドレスに変換して書き込み対象とするドラ

イブとドライブ内アドレスを決定する。制御プロセッサ119, 120は、共通制御メモリ112上のプロセッサ間連絡領域を参照し、データ読み出し要求があれば、該当するドライブ127～130からデータを読み出し、キャッシュメモリ122上の該当するキャッシュスロットに書き込み、キャッシュスロット管理テーブル上で該当するデータ有に更新する。

【0026】

ドライブ127～130への入出力要求は、制御プロセッサ119, 120のいずれでも処理可能である。例えば制御プロセッサ119またはファイバインタフェースのループ125が故障等で使用できないとき、制御プロセッサ120とファイバインタフェースのループ126によって入出力要求を処理できる。これにより障害発生時もドライブ127～130への入出力を停止することなく入出力要求を実行できる。

【0027】

制御プロセッサ114, 115, 116及び117は、お互いの制御プロセッサの状態を監視している。各制御プロセッサは、一定周期で共通制御メモリ112に現在時刻を書き込む。それを他の制御プロセッサが一定周期で参照し、前回参照時と時刻の差分があるか否かをチェックし、差分がなければ当該制御プロセッサは停止していると判断する。停止状態を見つけた制御プロセッサは、F C A L管理情報113からその制御プロセッサの管理情報を引き継ぎ、処理を継続する。例えば制御プロセッサ114が制御プロセッサ115の停止を検出したとする。このとき制御プロセッサ114は、F C A L管理情報113を図5のように書き換える。これにより制御プロセッサ114は、制御プロセッサ115が処理していたLUN10～19に関する入出力要求を引き継いで処理することが可能である。

【0028】

また制御プロセッサ114～117は、処理した入出力要求の処理数を計数し、一定周期で共通制御メモリ112に格納する。他の制御プロセッサは、この処理数を参照し、処理数の多い制御プロセッサを見つけ出し、より処理数の少ない制御プロセッサの処理数を上げることにより、各制御プロセッサの処理数を平均

化する。例えば制御プロセッサ117が制御プロセッサ116の処理数の低下と制御プロセッサ115の処理数の上昇を検出したとする。制御プロセッサ117は、F C A L管理情報113を図6のように書き換える。ただしループ133を介してA L - P AがE8を有するフレームが制御プロセッサ116に伝送されるように、P B C 108～111のうちの該当するP B Cのスイッチ制御を変更しなければならない。これにより制御プロセッサ116は、L U N 10～19及びL U N 20～29に関する入出力要求を処理することが可能となり、制御プロセッサ間の処理数を平均化し、負荷分散を実現できる。

【0029】

なお制御プロセッサの管理するL U N 203のうち一部のL U Nのみを他の制御プロセッサが引き継いで処理を継続することも可能である。例えば制御プロセッサ115の管理するL U N 10～19のうち制御プロセッサ116がL U N 15～19のみを引き継ぐことも可能である。このときF C A L管理情報113は図7のように書き換えられる。ただしL U N 15～19についてA L - P A 202とL U N 203との対応づけが変更となるので、制御プロセッサは、ホストコンピュータ100, 101, 102にこの変更を通知しなければならない。

【0030】

なお上記制御プロセッサ114～117の処理の流れは、ファイバチャネルインターフェースによってディスク制御装置107と接続されるホストコンピュータ100に関する入出力要求の処理について説明した。ホストコンピュータ101, 102は、ファイバチャネルインターフェースとは異なるインターフェースによってディスク制御装置107と接続するので、H I F C 104および105は、ホストコンピュータ101及び102から受け取った入出力コマンドをファイバチャネルインターフェースに従うフレームの形式に変換した後に、ループ133を介して制御プロセッサ114～117へ送信する。これによってホストコンピュータ101および102から送られる入出力要求の処理は上記処理と同様となる。H I F C 104は、E S C O N(Enterprise System Connection)と呼ばれるインターフェースに従うコマンド、制御情報、データとファイバチャネルインターフェースに従うコマンド、制御情報、データとの間の変換をする機能を有する。またH

I F C 1 0 5 は、 S C S I に従うコマンド、制御情報、データとファイバチャネルインタフェースに従うコマンド、制御情報、データとの間の変換をする機能を有する。このようにホストインタフェースの変換機能を有する H I F C をディスク制御装置 1 0 7 に搭載することによって、ホストインタフェースの種類に関係なく、任意のホストコンピュータをディスク制御装置 1 0 7 に接続することができる。

【0031】

また本実施形態はドライブ 1 2 7 ~ 1 3 0 がディスクドライブの場合について説明したが、 D I F C 1 2 3 , 1 2 4 を変更することによって磁気テープ装置やフロッピディスク・ドライブを接続することができる。また D I F C に、 S C S I とファイバチャネルインタフェースの変換機能を設けることによって、ループ 1 2 5 , 1 2 6 を S C S I によるケーブルに置き換えることもできる。

【0032】

本実施形態のディスク制御装置 1 0 7 によれば、ホストコンピュータ 1 0 0 から送られる入出力要求は制御プロセッサ 1 1 4 ~ 1 1 7 のうちのどの制御プロセッサでも処理できる。従ってホストコンピュータ 1 0 0 と H I F C 1 0 3 との間およびループ 1 3 3 のデータ転送速度に応じてホストコンピュータ 1 0 0 からの入出力要求の数が多い場合には、制御プロセッサ 1 1 4 ~ 1 1 7 すべてがホストコンピュータ 1 0 0 からの入出力要求を処理することが可能であり、制御プロセッサの台数がより少ない場合に比べてスループットが向上する。同様にホストコンピュータ 1 0 1 及び 1 0 2 から送られる入出力要求は、各々制御プロセッサ 1 1 4 ~ 1 1 7 のうちのどの制御プロセッサでも処理できる。このようにホストコンピュータ 1 0 0 , 1 0 1 及び 1 0 2 がループ 1 3 3 及び制御プロセッサ 1 1 4 ~ 1 1 7 を共有するので、従来のようにホストコンピュータ 1 0 0 , 1 0 1 及び 1 0 2 ごとに共通バスに接続するホストインタフェース部が独立している構成に比べて機構の分割損をなくし、記憶制御装置の性能向上を図るとともにコスト/性能比の向上を図ることができる。

【0033】

【発明の効果】

以上述べたように本発明によれば、ホストコンピュータから送られる入出力要求を複数の制御プロセッサによって並列処理するとともに、制御プロセッサ間で負荷の分散をするので、記憶制御装置の性能向上を図ることができる。特に高速のファイバチャネルの性能を充分生かすことができる。また1つの制御プロセッサが障害により停止した場合に、他の制御プロセッサが障害制御プロセッサの処理を引き継ぐので、信頼性の高い記憶制御装置を提供できる。

【0034】

さらに本発明の記憶制御装置は、複数種類のインターフェースをもつホストコンピュータを接続可能であるとともに、複数のホストコンピュータが記憶制御装置内部のファイバチャネルループ及び制御プロセッサを共有するので、コスト／性能比のよい記憶制御装置を提供できる。また異なる種類の記憶媒体の駆動装置を接続可能な記憶制御装置を提供できる。

【図面の簡単な説明】

【図1】

実施形態の記憶サブシステムの構成図である。

【図2】

実施形態のループ133と関連機構の構成図である。

【図3】

実施形態のF C A L管理情報113のデータ構成を示す図である。

【図4】

実施形態の制御プロセッサ114～117の処理の流れを示すフローチャートである。

【図5】

制御プロセッサが停止したときに書き換えたF C A L管理情報113の例を示す図である。

【図6】

制御プロセッサの負荷の不均衡を検出したときに書き換えたF C A L管理情報

113の例を示す図である。

【図7】

制御プロセッサの負荷の不均衡を検出したときに書き換えた他のF C A L 管理情報113の例を示す図である。

【図8】

従来の記憶制御装置の構成例を示す図である。

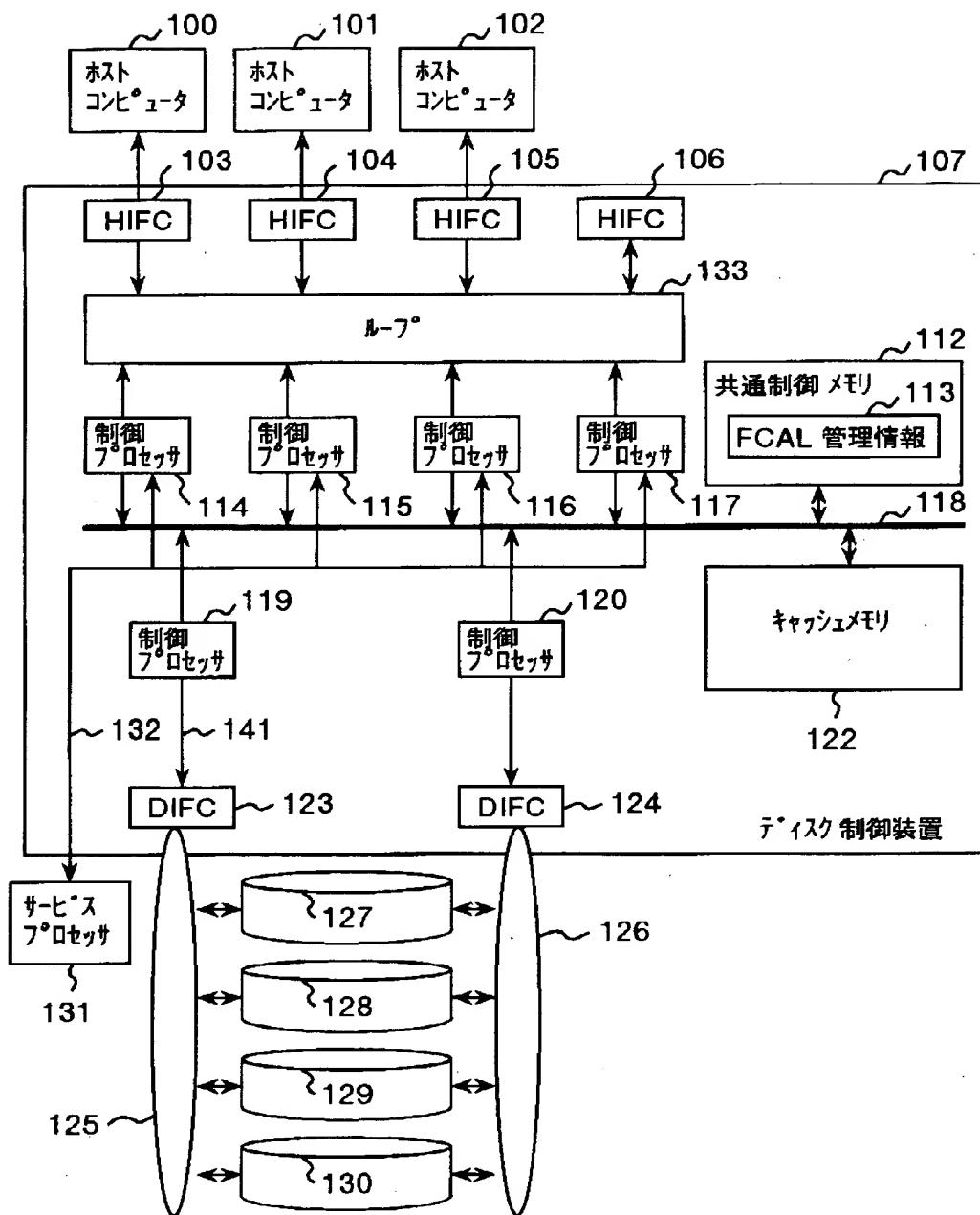
【符号の説明】

100, 101, 102: ホストコンピュータ、103, 104, 105, 106: ホストインターフェースコントローラ、107: ディスク制御装置、113: F C A L 管理情報、114~117, 119~120: 制御プロセッサ、123, 124: ドライブインターフェースコントローラ、127, 128, 129, 130: ドライブ

【書類名】 図面

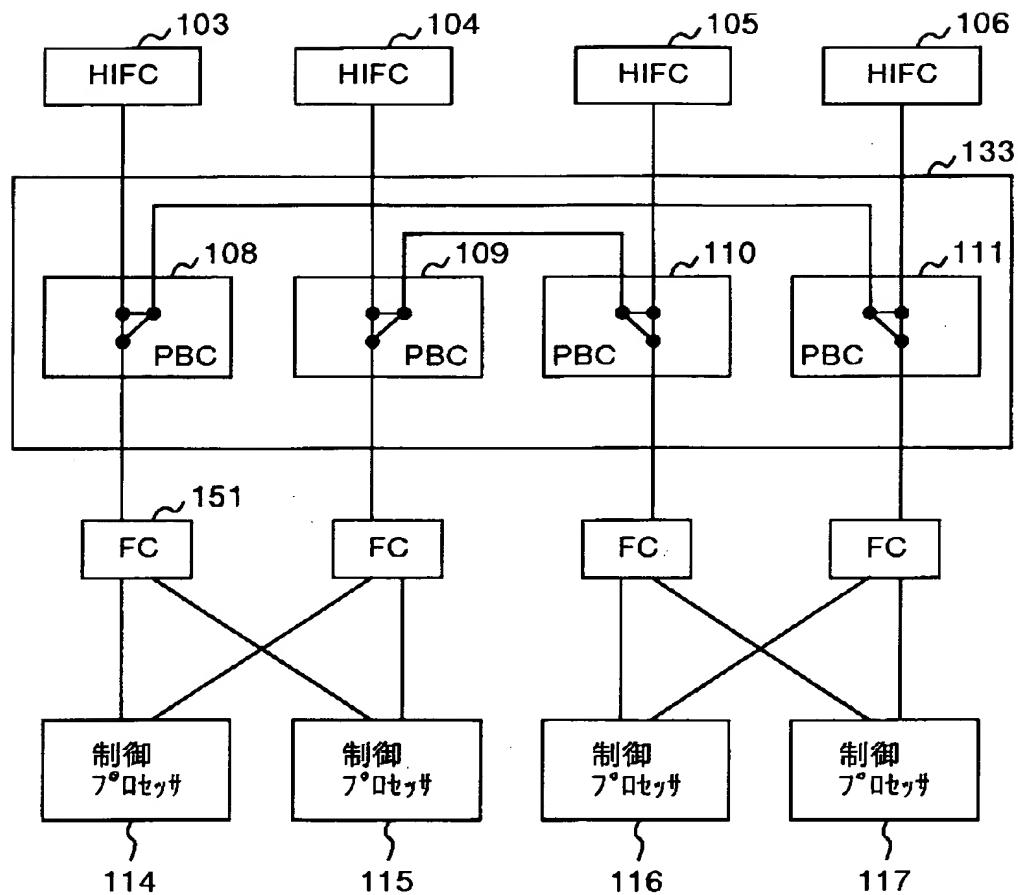
【図1】

図 1



【図2】

図 2



【図3】

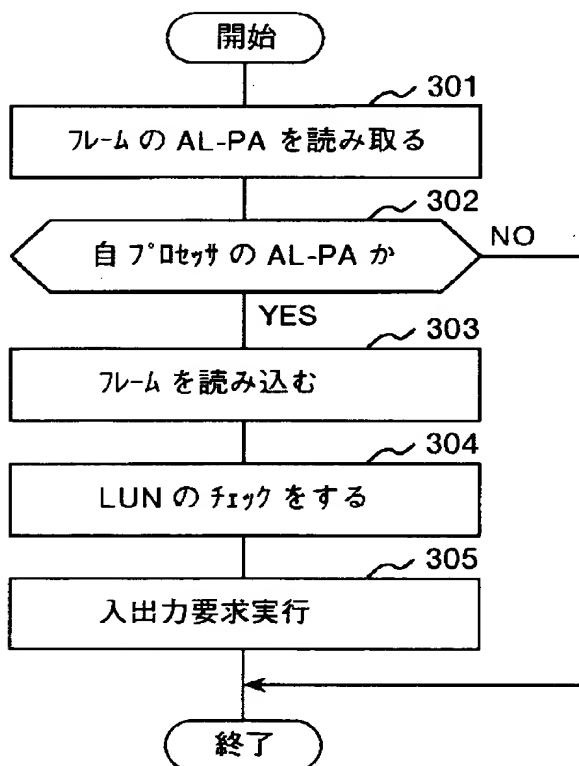
図3

113 : F C A L 管理情報

制御プロセッサ番号	AL-PA	LUN
114	E F	0-9
115	E 8	10-19
116	E 4	20-29
117	E 2	30-39

【図4】

図4



【図5】

図5

113 : F C A L 管理情報

制御プロセッサ番号	201 AL-PA	202	203 LUN
114	E F		0-9
114	E 8		10-19
115	-		-
116	E 4		20-29
117	E 2		30-39

【図6】

図6

113 : F C A L 管理情報

制御プロセッサ番号	201 AL-PA	202	203 LUN
114	E F		0-9
115	-		-
116	E 8		10-19
116	E 4		20-29
117	E 2		30-39

【図7】

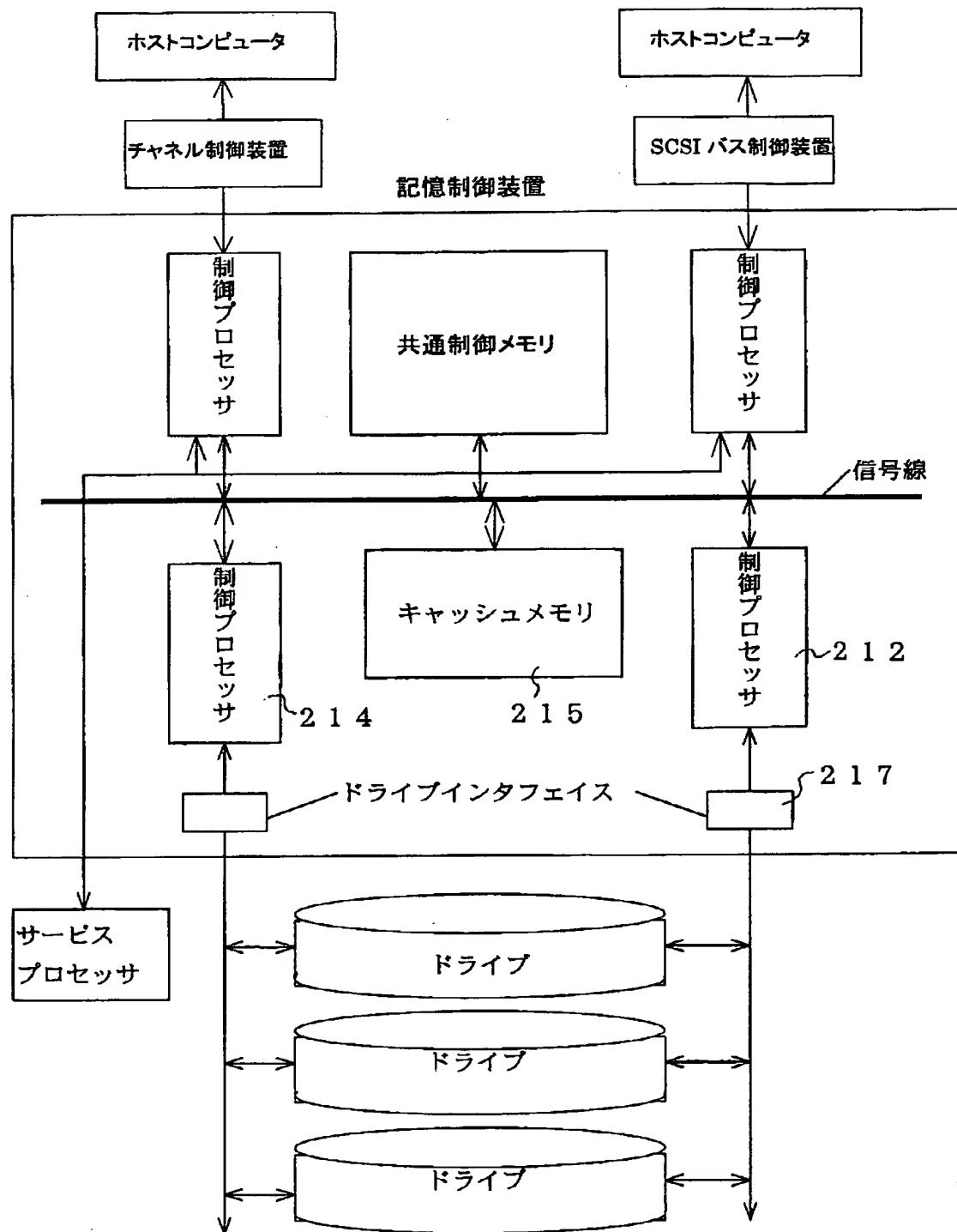
図7

113 : F C A L 管理情報

制御プロセッサ番号	201 AL-PA	202	203 LUN
114	E F		0-9
115	E 8		10-14
116	E 4		15-29
117	E 2		30-39

【図8】

図8



【書類名】 要約書

【要約】

【課題】 記憶制御装置の性能向上、特にファイバチャネルのもつ高速データ転送を生かすような高性能をもつとともに、信頼性の向上を図る。また複数種類のインターフェースをもつホストコンピュータを接続可能とする。

【解決手段】 ループ133はファイバチャネルインターフェースをもつ共通のループ伝送路である。HIFC103, 104, 105は、各々インターフェースの異なるホストコンピュータ100, 101, 102と接続し、必要に応じファイバチャネルインターフェースとの間の変換機能を有する。制御プロセッサ114～117は、HIFC103～105によって共有されるプロセッサである。制御プロセッサ114～117は、各々FCAL管理情報113を参照し、ループ133を流れるフレームのうち設定されたアドレスをもつフレームを取り込み、設定されたLUN範囲の入出力要求の処理をする。

【選択図】 図1

出願人履歴情報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所

For: The Patent Application

Our case No. NT0083US

THE PRIOR ART REFERENCES CITED IN THE SPECIFICATION

1. Japanese Unexamined Patent Publication No. Hei 05-143242



Home

Search

List

Include

MicroPatent® PatSearch FullText: Record 1 of 1

Search scope: US; EP; WO; JP**Years:** 1976-2000 **Patent/Publication No.:** JP05143242**[no drawing available]**[Download This Patent](#)[Family Lookup](#)[Go to first matching text](#)**JP05143242****MAGNETIC DISK SYSTEM****HITACHI LTD****Inventor(s):** ARASAWA NOBUYUKI ; TOYODA MITSURU ; OYAMA MITSUO
Application No. 03308829, **Filed** 19911125, **Published** 19930611**Abstract:** PURPOSE: To constitute the system so that a host computer can execute an access to the magnetic disk system by setting an optimal function to every channel.**CONSTITUTION:** Between a disk controller 2 and a disk device group 12, a cache circuit 14 and a data retrieving circuit 15, a switching means for switching arbitrarily a connecting relation of outputs of the cache circuit 14 and the data retrieving circuit 15, and a disk controller side bus 5, and a switching means for switching arbitrarily inputs of the cache circuit 14 and the data retrieving circuit 15, and a data bus 13 of the disk device group are provided.**COPYRIGHT:** (C)1993,JPO&Japio**Int'l Class:** G06F00306; G06F01314**MicroPatent Reference Number:** 000196526**COPYRIGHT:** (C)JPO

Home

Search

List

Include

For further information, please contact:

[Technical Support](#) | [Billing](#) | [Sales](#) | [General Information](#)

(19)日本国特許庁 (JP)	(12)公開特許公報 (A)	(11)特許出願公開番号	特開平5-143242
(22)出願日	平成5年(1993)6月11日	技術表示箇所	審査請求 未請求 審査項の数4(全16頁)
(21)出願番号	特願平3-308829	(71)出願人	株式会社日立製作所 000005108
(22)出願日	平成3年(1991)11月25日	(72)発明者	荒澤 伸幸 東京都千代田区神田駿河台四丁目6番地 株式会社日立製作所中央研究所内
		(72)発明者	豊田 誠 東京都千代田区神田駿河台1丁目280番地 株式会社日立製作所中央研究所内
		(72)発明者	大山 光男 東京都千代田区神田駿河台1丁目280番地 株式会社日立製作所中央研究所内
		(74)代理人	弁理士 小川 駿男

【特許請求の範囲】
 【請求項1】複数のチャネルを持つホスト計算機の制御下にあり、複数のデータバスを制御するディスク制御装置と、前記ディスク制御装置の制御下に、前記複数のデータバスを持つディスク装置群とからなる磁気ディスクシステムにおいて、前記ディスク装置群との間に、前記ディスク装置群からのデータ読み出しを高速化するデータ検索処理手段と、前記データ読み出し手段と前記ディスク装置群との間に、前記ディスク装置群からのデータ検索を行なうデータ検索処理手段と、前記二つの処理手段の出力と前記ディスク制御装置のデータバスの接続關係を任意に切り換えることのできる、ディスク制御装置側切り換え手段と、前記二つの処理手段の入力と前記ディスク装置群からのデータ読み出し手段との間に、前記ディスク装置群からのデータ読み出しを高速化するデータ検索処理手段と、前記データ読み出し手段と前記ディスク装置群との間に、前記ディスク装置群からのデータ読み出しを高速化するデータ検索処理手段と、前記二つの処理手段の出力と前記ディスク装置群とのデータバスの接続關係を任意に切り換えることのできるディスク制御装置側切り換え手段と、前記二つの処理手段を設けたことを特徴とする磁気ディスクシステム。

【請求項2】複数のチャネルを持つホスト計算機の制御下にあり、複数のデータバスを制御するディスク制御装置と、前記ディスク制御装置の制御下にあり、前記複数のデータバスを持つディスク装置群とからなる磁気ディスクシステムにおいて、前記ディスク制御装置と前記ディスク装置群との間に、前記ディスク装置群からのデータ読み出しを高速化する高圧縮出し処理手段と、前記高圧縮出し手段と同一機能をもつ予備の高圧縮出し処理手段と、前記高圧縮出し手段と、前記二つの処理手段の出力と前記ディスク制御装置のデータバスの接続關係を任意に切り換えることのできるディスク制御装置側切り換え手段と、前記二つの処理手段の入力と前記ディスク装置群とのデータバスの接続關係を任意に切り換えることのできるディスク制御装置側切り換え手段を設けたことを特徴とする磁気ディスクシステム。

【請求項3】複数のチャネルを持つホスト計算機の制御下にあり、複数のデータバスを制御するディスク制御装置と、前記ディスク制御装置の制御下にあり、前記複数のデータバスを持つディスク装置群とからなる磁気ディスクシステムにおいて、前記ディスク制御装置と前記ディスク装置群との間に、データの検索を行なう複数個のデータ検索処理手段と、前記ディスク制御装置側切り換え手段と前記複数のデータ検索処理手段を任意に切り換えることのできるディスク制御装置側切り換え手段と前記複数のデータ検索処理手段の出力を任意に切り換えることのできるディスク制御装置側切り換え手段を設けたことを特徴とする磁気ディスクシステム。

【請求項4】複数のチャネルを持つホスト計算機の制御下にあり、複数のデータバスを制御するディスク制御装置と、前記ディスク制御装置の制御下にあり、前記複数のデータバスを持ち、読み出したデータを、同時に前記複数のデータバスに出力することのできる、ディスク装置群とからなる磁気ディスクシステムにおいて、前記データバスに接続する部品(例えはデータ検索など)に接続して性能を向上させることのできない(ステップ1114)。

(57)【要約】	【構成】ディスク制御装置2と、ディスク装置群1,2と、データバス回路1,4とデータ検索回路1,5、キャッシュ回路1,4とデータ検索回路1,5との接続關係を任意に切り換える手段と、キャッシュ回路1,4とデータ検索回路1,5の入力とデータバスの接続關係を任意に切り換える手段と、キャッシュ回路1,4とデータ検索回路1,5の入力とディスク装置群のデータバス1,3の接続關係を任意に切り換える手段を設ける。	【効果】ホスト計算機は、チャネルごとに最適な機能を設定して磁気ディスクシステムにアクセスできる。
(58)【発明の名称】磁気ディスクシステム	図1	図1
(59)【発明の詳細な説明】 【発明の詳細な説明】本発明は、複数のチャネルに接続される磁気ディスクシステムに係り、特に、各チャネル間の接続關係を任意に切り換える手段と、前記データ検索処理手段を設けたことを特徴とする磁気ディスクシステム。	【発明の詳細な説明】 【発明の利用分野】本発明は、複数のチャネルに接続される磁気ディスクシステムに係り、特に、各チャネル間の接続關係を任意に切り換える手段と、前記データ検索処理手段を設けたことを特徴とする磁気ディスクシステム。	
(60)【発明の図】 <img alt="Block diagram of the magnetic disk system (Figure 1). The diagram shows a host computer (ホスト計算機) connected to a disk controller (ディスク制御装置) via bus lines (バス回線). The disk controller is connected to two disk drives (ディスク装置群) labeled 1 and 2. Each disk drive has its own cache memory (キャッシュ回路) and a data search circuit (データ検索回路). The disk controller also has four bus lines (バス回線) labeled 1, 4, 5, and 10. Bus line 1 connects to bus line 1 of each disk drive. Bus line 4 connects to bus line 4 of each disk drive. Bus line 5 connects to bus line 5 of each disk drive. Bus line 10 connects to both disk drives. There are switches (スイッチ) labeled 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 778, 779, 779, 780, 781, 782, 783, 784, 785, 786, 787, 787, 788, 789, 789, 790, 791, 792, 793, 794, 795, 796, 797, 797, 798, 799, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 878, 879, 879, 880, 881, 882, 883, 884, 885, 886, 887, 887, 888, 889, 889, 890, 891, 892, 893, 894, 895, 895, 896, 897, 897, 898, 898, 899, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 978, 979, 979, 980, 981, 982, 983, 984, 985, 986, 987, 987, 988, 989, 989, 990, 991, 992, 993, 994, 995, 995, 996, 997, 997, 998, 998, 999, 999, 1000, 1001, 1002, 1003, 1004, 1005, 1006, 1007, 1008, 1009, 1009, 1010, 1011, 1012, 1013, 1014, 1015, 1016, 1017, 1018, 1019, 1019, 1020, 1021, 1022, 1023, 1024, 1025, 1026, 1027, 1028, 1029, 1029, 1030, 1031, 1032, 1033, 1034, 1035, 1036, 1037, 1038, 1039, 1039, 1040, 1041, 1042, 1043, 1044, 1045, 1046, 1047, 1048, 1049, 1049, 1050, 1051, 1052, 1053, 1054, 1055, 1056, 1057, 1058, 1059, 1059, 1060, 1061, 1062, 1063, 1064, 1065, 1066, 1067, 1068, 1069, 1069, 1070, 1071, 1072, 1073, 1074, 1075, 1076, 1077, 1078, 1078, 1079, 1079, 1080, 1081, 1082, 1083, 1084, 1085, 1086, 1087, 1087, 1088, 1089, 1089, 1090, 1091, 1092, 1093, 1094, 1095, 1095, 1096, 1097, 1097, 1098, 1098, 1099, 1099, 1100, 1101, 1102, 1103, 1104, 1105, 1106, 1107, 1108, 1108, 1109, 1109, 1110, 1111, 1112, 1113, 1113, 1114, 1114, 1115, 1116, 1117, 1118, 1118, 1119, 1119, 1120, 1121, 1122, 1123, 1123, 1124, 1124, 1125, 1126, 1126, 1127, 1127, 1128, 1128, 1129, 1129, 1130, 1130, 1131, 1131, 1132, 1132, 1133, 1133, 1134, 1134, 1135, 1135, 1136, 1136, 1137, 1137, 1138, 1138, 1139, 1139, 1140, 1140, 1141, 1141, 1142, 1142, 1143, 1143, 1144, 1144, 1145, 1145, 1146, 1146, 1147, 1147, 1148, 1148, 1149, 1149, 1150, 1150, 1151, 1151, 1152, 1152, 1153, 1153, 1154, 1154, 1155, 1155, 1156, 1156, 1157, 1157, 1158, 1158, 1159, 1159, 1160, 1160, 1161, 1161, 1162, 1162, 1163, 1163, 1164, 1164, 1165, 1165, 1166, 1166, 1167, 1167, 1168, 1168, 1169, 1169, 1170, 1170, 1171, 1171, 1172, 1172, 1173, 1173, 1174, 1174, 1175, 1175, 1176, 1176, 1177, 1177, 1178, 1178, 1179, 1179, 1180, 1180, 1181, 1181, 1182, 1182, 1183, 1183, 1184, 1184, 1185, 1185, 1186, 1186, 1187, 1187, 1188, 1188, 1189, 1189, 1190, 1190, 1191, 1191, 1192, 1192, 1193, 1193, 1194, 1194, 1195, 1195, 1196, 1196, 1197, 1197, 1198, 1198, 1199, 1199, 1200, 1200, 1201, 1201, 1202, 1202, 1203, 1203, 1204, 1204, 1205, 1205, 1206, 1206, 1207, 1207, 1208, 1208, 1209, 1209, 1210, 1210, 1211, 1211, 1212, 1212, 1213, 1213, 1214, 1214, 1215, 1215, 1216, 1216, 1217, 1217, 1218, 1218, 1219, 1219, 1220, 1220, 1221, 1221, 1222, 1222, 1223, 1223, 1224, 1224, 1225, 1225, 1226, 1226, 1227, 1227, 1228, 1228, 1229, 1229, 1230, 1230, 1231, 1231, 1232, 1232, 1233, 1233, 1234, 1234, 1235, 1235, 1236, 1236, 1237, 1237, 1238, 1238, 1239, 1239, 1240, 1240, 1241, 1241, 1242, 1242, 1243, 1243, 1244, 1244, 1245, 1245, 1246, 1246, 1247, 1247, 1248, 1248, 1249, 1249, 1250, 1250, 1251, 1251, 1252, 1252, 1253, 1253, 1254, 1254, 1255, 1255, 1256, 1256, 1257, 1257, 1258, 1258, 1259, 1259, 1260, 1260, 1261, 1261, 1262, 1262, 1263, 1263, 1264, 1264, 1265, 1265, 1266, 1266, 1267, 1267, 1268, 1268, 1269, 1269, 1270, 1270, 1271, 1271, 1272, 1272, 1273, 1273, 1274, 1274, 1275, 1275, 1276, 1276, 1277, 1277, 1278, 1278, 1279, 1279, 1280, 1280, 1281, 1281, 1282, 1282, 1283, 1283, 1284, 1284, 1285, 1285, 1286, 1286, 1287, 1287, 1288, 1288, 1289, 1289, 1290, 1290, 1291, 1291, 1292, 1292, 1293, 1293, 1294, 1294, 1295, 1295, 1296, 1296, 1297, 1297, 1298, 1298, 1299, 1299, 1300, 1300, 1301, 1301, 1302, 1302, 1303, 1303, 1304, 1304, 1305, 1305, 1306, 1306, 1307, 1307, 1308, 1308, 1309, 1309, 1310, 1310, 1311, 1311, 1312, 1312, 1313, 1313, 1314, 1314, 1315, 1315, 1316, 1316, 1317, 1317, 1318, 1318,		

[0005] このように従来の技術は、特定の機能を磁気ディスクシステムに付加するためには有効であるが、一台のディスクシステムに複数の性能、あるいは機能、例えば、キャッシュを用いた高速読み出し機能とデータ検索の高い機能の二つの機能を持たせ、かつ必要に応じて使い分けることができる磁気ディスクシステムについても考案されていなかった。このため、このような場合は回遊されなければならない。

[0006] また既存技術では、性能の向上のための手段、例えば、図9に示すようなデータバッファが、ディスク装置群のバスに固定して設けられているので、バッファメモリが故障した場合には、ディスク装置群のデータがアクセス不能となる場合があり、コンピュータシステムに大きな影響を与える。

[0007] また、検索するデータ量が増加しているため、データ検索時間に占める磁気ディスクシステムからホスト計算機への、データ伝送時間の割合が大きくなってしまっており、データ伝送時間の短縮が望まれている。

[0008] 本発明の目的は、一台のディスクシステムに複数の機能、性能を持たせ、必要に応じて軟件的に使い分けができるような磁気ディスクシステムを実現することにある。

[0009] 本発明の他の目的は、性能向上の手段、新しい機能の付加手段に障害の発生した場合にも、障害部分を回避することにより、障害がシステムに重大な影響を及ぼさない磁気ディスクシステムを実現することにある。

【0010】 離題を解決するための手段 上記の目的を達成するた
めに、複数のチャネルを介してホスト計算機に接続され
るディスク制御装置と、複数のデータバスを持つディス
ク装置群からなる磁気ディスクシステムにおいて、前記
ディスク制御回路と前記ディスク装置群との間に、前記
ディスク装置群からのデータ読み出しを高速化させるた
めのチャネルと、読み出したデータを検査して、必要
なデータのみをホスト計算機に送るための検索手段と、
前記データ検索手段の出力を前記データ検索手段と
前記チャネルと、前記チャネルの出力を前記データ検
索手段の入力を、前記ディスク装置群のバスと接
続するための下位スイッチと、二つのスイッチを前記デ
ィスク装置群からの信号により開閉するスイッチ制御回
路を設け、ホスト計算機で実行されるジョブが必要と
する機能を選択できるようにした。
【0011】 さらに、チャッシュを、少なくとも一個の
代群のチャッシュを含めて構成し、前記チャッシュに附
着が発生した場合は、階層が発生したチャッシュを代替
チャッシュに切り換えるようにした。
【0012】 さらに、各個のデータチャネルを介して、か

[001.3] さらに、複数のデータバスを持つディスク装置に、複数のデータバスに同一データを出力する手段として、複数のデータ検索手段を駆け、検索するデータを複数のデータ検索手段に同時に投入して検索能力を高めるようにした。

[0014] 段、例えば、図9に示すようなデータバッファが、ディスク装置群のバスに固定して設けられてるので、バスファミモリが故障した場合には、ディスク装置群のデータアクセス不能となる場合があり、コンピュータシステムに大きな影響を与える。

[0017] また、検索するデータ量が増加しているため、データ検索時間に占める磁気ディスクシステムからホスト計算機への、データ転送時間の割合が大きくなってしまっており、データ転送時間の短縮が望まれている。

[0018] 本発明の目的は、一台のディスクシステムに複数の機能、性能を持たせ、必要に応じて柔軟に使い分けができるような磁気ディスクシステムを実現することである。

[0019] 本発明の他の目的は、性能向上の手段、新しい機能の付加手段による障害の発生した場合には、障害部分を回避することにより、障害がシステムに重大な影響を及ぼさない磁気ディスクシステムを実現することである。

[0020] 本発明では、図12に示すように、ディスク制御装置が、スイッチ制御回路に切り換えるタイミングを制御して、データ検索の場合には、キャッシュ回路を選択し、高速データ転送が必要な場合には、キャッシュ回路を選択してできる。すなわち、ホスト計算機が必要とするアクセス特性に合わせて、一台の磁気ディスクシステムを有する磁気ディスクシステムと使用でき、効率の良いデータアクセスを実現することができる。また、キャッシュ回路とデータ検索回路は複数のチャネルに対応して並列に動作可能な構造であり、あたかも、黒い黒い機能を持つ複数台のディスクシステムのように動作可能となる。

[0021] また、キャッシュ回路とデータ検索回路を、キャッシュ回路と代替キャッシュ回路とで構成することにより、キャッシュ回路に障害が生じたときは代替キャッシュ回路に切り換えて、障害によるディスクシステムに大きな影響を与えない。

たところに特徴がある。

【0022】ホスト計算機1にディスク装置1.2(1)からデータを読み込む場合、ホスト計算機1は、チャネル4(1)～4(4)のうち、空きのチャネルを用いて、ディスク制御装置2にデータ読み込みを行いうりで、データ読み込みを行いうりで、ディスク制御装置2はバス5コマンドを発行する。ディスク制御装置2はバス5(1)～5(4)、キャッシュ回路1.4(1)～1.4(2)、データバス1.3(1)～1.3(4)を調べる。この時、例えバス5(1)とキャッシュ回路1.4(1)とデータバス1.3(1)が空いていた場合、ディ

キヤッシュにヒットした場合のデータ読み込み時間を早くするキャッシュ回路、1.5(1)～1.5(2)は読み出しだデータの検索を行なうデータ検索回路、7はバス5(1)～5(4)をキャッシュ回路1.4(1)～1.4(2)の出力、およびデータ検索回路1.5(1)～1.5(2)の出力に任意に接続する上位スイッチ、8はキャッシュ回路1.4(1)～1.4(2)の入力、およびデータ検索回路1.5(1)～1.5(2)の入力とディスク装置1.2(1)～1.2(n)とを、データバス1.3(1)～1.3(4)を介して、任意に接続する下位スイッチ、

9はディスク制御装置2から出力されるスイッチ制御コマンド6を経由して、上位スイッチ7と下位スイッチ8を切り換えて、バス5(1)とキャッシュ回路14(1)とデータバス1.3(1)を接続する。

[0019] 図2において、20はディスク制御装置2から発行されたコマンドを入力するコマンドバッファ、21はコマンドバッファに入力されたコマンドを解読するコマンドデコーダ、22はコマンドデコードで解読された内容がキャッシュに関係したコマンドであつたときにキャッシュを削除するキャッシュ制御回路、25はキャッシュから読み出されたデータを保持する出力バッファア、23はディスク装置1.2(1)～1.2(2)から読み出したデータを保持する入力バッファ、24は入力バッファ2の内容をキャッシュモモリ、2.6は出力バッファ2.5のバッファ(データ書き込み時にディスク制御装置2によって付加)チェックを行なうパリティエラーチェック回路、2.7は出力バッファア(パリティエラーが発生した時に出力される出力エラー信号である。

[0020] 図3において、30はディスク制御装置2から発行されたコマンドを入力するコマンドバッファ、31はコマンドバッファに入力されたコマンドを解読するコマンドデコーダ、32はコマンドデコードであつたときのデータがデータ検索に関係したコマンドであつたときに比較器3.5を制御する比較器制御回路、3.3はディスク装置1.2(1)～1.2(2)から読み出したデータを保持する入力バッファ、3.4はディスク制御装置2から送られてきた比較データを保持する比較データレジスタ、3.6是比较データレジスタ3.4の内容と入力バッファ3.3の内容の条件が一致したときにそのデータを保持する一致データレジスタ、3.7は出力バッファ3.6のパリティ(データ書き込み時にディスク制御装置2によつて付加)チェックを行なうパリティエラーチェック回路、3.8は出力バッファア(パリティエラーが発生した時に出力される出力エラー信号である。

[0021] 本実施例では、ホスト計算機1へのデータの読み込みと検索が、頻繁に発生するような磁気ディスクシステムにおいて、キャッシュ回路1.4(1)～1.4(2)とデータ検索回路1.5(1)～1.5(2)と下位スイッチ7とデータバス5(1)とデータバス1.3(1)が行なう磁気ディスクシステムとデータ検索回路1.5(1)～1.5(2)とデータバス5(1)とデータバス1.3(1)が行なう場合、ディスク制御装置2はバス5(1)～5(4)、データ検索回路2はバス5(1)～5(4)、データバス1.3(1)～1.3(4)を調べる。この時、例えばバス5(1)とデータ検索回路1.5(1)とデータバス1.3(1)が行なう場合、ディスク制御装置2は

は、スイッチ制御回路 9 にスイッチ制御コマンド 6 を発行して、上位スイッチ 7 と下位スイッチ 8 を切り換えて、バス 5 (1) とデータ検索回路 15 (1) とデータを繋続する。更にデータ検索回路 15 (1) にデータを転送して、データを読み込みだ時の動作を設定する。この時、データ検索回路 15 (1) ではコマンドバッファ 3 0 にコマンドを読み取り込み、コマンドコードアダ 3 1 で解読を行い、比較器 4 (1) ～ 4 (4) のうち、空きのチャネルを用いて、データを転送する。ディスク制御装置 1 はバス 5 (1) ～ 4 (4) に検索データを書き込む。

もしキャッシュ回路 14 (1) ~ 14 (4) のうち、どれかが故障した場合、予備キャッシュ回路 14 (5) に切り換える事により、磁気ディスクシステムが性能低下下、あるいは使用不能になることなく処理を続けることができる。このように、第一の実施例では、耐障害性の高い磁気ディスクシステムを実現している。

[0036] 図 5 は本発明の第三の実施例を示す図である。図 6 はバス出力スイッチ 51 の構成図である。

[0037] 図 5において、15 (1) ~ 15 (4) は読み込んだデータの検索を行うデータ検索回路、51

〔0038〕本実施例では、磁気ディスクシステムにおいて、読み出したデータに対し、データ検索を行って、必要なデータのみをホスト計算機1に転送する場合、ディスク基盤12(2)から読み込んだデータを全バスに出力する全バス出力サイクルを実現する。本実施例の磁気ディスクシステムにおいて、同一データに対し、データ検索を行った結果を、ホスト計算機1に読み込む場合、ホスト計算機1は、チャネル4(1)～4(4)のうち空きのチャネルを用いて、ディスク制御装置2に検索コマンドと三種類の異なる検索データを転送する。ディスク制御装置2は検索データを複数のデータ検索回路1～15(4)とディベリバ

【0039】 それぞれのデータ検索回路 1.5 (1) ~ 1.5 (3) に同一データが転送される。
【0040】 ディスク制御装置 2 は、このデータをチャネル 3.4 内にある比較データと比較し、比較データレジスタ 3.4 内にある比較データと比較器 3.5 によって比較を行ない、条件を満たしたデータだけを一数データレジスタ 3.6 に書き込む。この後三つのデータ検索回路 1.5 (1) ~ 1.5 (3) 内の一数データをディスク制御装置 2 に転送する。

【004-1】 図7は本発明の第四の実施例を示す図である。図8は図7における金バスディスク装置71(1)～71(n)の構成図である。

【004-2】 図7において、1.5(1)～1.5(4)は読み込んだデータの検索を行なうデータ検索回路、7.1(1)～7.1(2)は読み出したデータを任意のバス1～3(1)～1.3(4)に出力することができる金バスディスク基盤である。

【004-3】 本実施例では、磁気ディスクシステムにおいて、読み出したデータに対して、データ検索を行ない、必要なデータのみをホスト計算機1に転送する場合、ディスク装置71(2)から読み出したデータを金バスディスク装置71(1)～7.1(2)を用いて、複数のデータ検索回路1.5(1)～1.5(4)に同一データを入力して、データ検索処理を並列に行なう磁気ディスクシステムを実現するところに特徴がある。

【004-4】 本実施例の磁気ディスクシステムにおいて、同一データについて三種類のデータ検索を行なつたデータを転送する。ディスク制御装置2はバス5(1)～5(4)、データバス1.3(1)～1.3(4)を調べる。この時、例えばバス5(1)と三つのデータ検索回路1.5(1)～1.5(3)と三つのデータバス1.3(1)～1.3(3)が空いた場合、ディスク制御装置2は、1.3(1)と3つのデータ検索回路1.5(1)～1.5(3)と三つのデータバス1.3(1)～1.3(3)を接続する。

る。更に三つのデータ検索回路 1.5 (1) ~ 1.5 (3) に制御コマンドと三つの異なる検索データを転送して、データを読み込んだ時の動作を観察する。この時、三つ のデータ検索回路 1.5 (1) ~ 1.5 (3) ではコマンドバッファ 3.0 にコマンドを取り込み、コマンドコード 3.1 で解説を行い、比較器制御回路 3.2 の動作を観察し、比較データレジスタ 3.3 に比較データを書き込む。

この後、ディスク制御装置 2 は全バスディスク装置 7.1 (2) にリードコマンドを発行する。全バスディスク装置 7.1 (2) は、データを読み出す。この時、全バスディスク装置 7.1 (2) はディスク制御装置 2 に指定されたデータバス 1.3 (1) ~ 1.3 (3) 全てに同一データを出力して、データ検索回路 1.5 (1) ~ 1.5 (3) にデータを転送する。それぞれのデータ検索回路 1.5 (1) ~ 1.5 (3) 内では入力バッファ 3.3 にデータを一時取り込み、比較データレジスタ 3.4 内にある比較データだけを一致データレジスタ 3.6 に書き込む。この後三つのデータ検索回路 1.5 (1) ~ 1.5 (3) 内の一組データをディスク制御装置 2 に伝送する。ディスク制御装置 2 は、このデータをチャネル 4 (1) ~ 4 (4) のうち空きのチャネルを使ってホスト計算機 1 にデータを伝送する。従来はデータ全てをホスト計算機 1 に読み込んでから検索を行なっていたが、データ検索回路 1.5 (1) ~ 1.5 (4) を設けることにより、不要なデータを取り除いて、必要なデータだけを送ることでデータを転送するため、転送データ量が少なく、データ伝送時間が短縮される。更に、本実施例では 3 つの検索を同時に行なうため、検索時間も短縮される。

【0045】図 1.1 は、本発明の実施例の効果を説明するフローチャートである。ホスト計算機からある処理が発行された場合 (ステップ 1.2.1, 1.2.2)、ディスク制御装置において処理の種類を調べる (ステップ 1.2.3)。リードの場合は、更にキャッシュ回路に故障が無いかチェックし (ステップ 1.2.4)、故障の無い場合はスイッチを切り換えてキャッシュ回路を選択し (ステップ 1.2.5)、処理を実行する。この時、キャッシュ回路はリード性能を向上させる (ステップ 1.2.7)。もし、故障があった場合は、予備のキャッシュ回路に切り換えて (ステップ 1.2.8)、同様にリード性能を向上させる (ステップ 1.2.9)。ディスク制御装置において処理の種類を調べた時 (ステップ 1.2.3)、データ検索の場合、データ検索回路に故障が無いかどうかをチェックし (ステップ 1.2.5)、無い場合はスイッチを切り換えてデータ検索回路を選択し (ステップ 1.3.0)、処理を実行する。この時データ検索回路は検索性能を向上させる (ステップ 1.3.1)。もし、故障があつた場合は、予備

のデータ検索回路に切り換えて (ステップ 1.3.2)、検索性能を向上することができる (ステップ 1.3.3)。

【0046】他の効果として、ディスク装置から読み出したデータを複数のデータ検索回路に入力してデータ検索処理を並列に行なうことにより、データ検索を高速に行う磁気ディスクシステムを実現することもできる。

【0047】【発明の効果】本発明では、ホスト計算機へのデータの読み込みと検索が、順繋りに動作するような磁気ディスクシステムにおいて、キャッシュ回路とデータ検索回路をスイッチを用いて切り換える、キャッシュ回路を使用するにより、キャッシュにヒットした場合は、ディスクの回転待ちナーシーク時間分早く終わることができるので、特にシーケンシャルリードの場合は特にヒットの確率が高く、高速にデータ伝送を行うことができる。データ検索の場合は、データ検索回路により、不要データをホスト計算機に送らないので、ホスト計算機へのデータ伝送時間が短縮される。すなわち、本発明によれば、チャネルごとに、複数の異なる性能、あるいは機能から、必要となる性能、あるいは機能を選択して同時に使用できるので、またかも、異なる性能を持つ複数台のディスクシステムのように動作可能となり、磁気ディスクシステムにおいて、効率のよいアクセスを行なうことができるので、上記の回路をキャッシュ回路と予備のキャッシュ回路で構成することにより、キャッシュ回路に障害が発生した場合にも、性能低下、あるいは使用不能になることなく処理を繰り返し、耐障害性の高い磁気ディスクシステムを実現することができる。

【図面の簡単な説明】

【図 1】本発明の第一の実施例を示すブロック図。

【図 2】キャッシュ回路のブロック図。

【図 3】データ検索回路のブロック図。

【図 4】本発明の第二の実施例を示すブロック図。

【図 5】本発明の第三の実施例を示すブロック図。

【図 6】全バス出力スイッチのブロック図。

【図 7】本発明の第四の実施例を示すブロック図。

【図 8】全バスディスク装置のブロック図。

【図 9】従来例の説明図。

【図 10】従来例の効果と問題点を表すフローチャート。

【図 11】本発明の効果を説明するフローチャート。

【図 12】本発明の作用を説明するタイムチャート。

【符号の説明】
1…ホスト計算機、2…ディスク制御装置、7…上位スイッチ、8…下位スイッチ、9…スイッチ制御回路、5…全バス出力スイッチ、7.1 (1) ~ 7.1 (n)…全バスディスク。

[図 1]

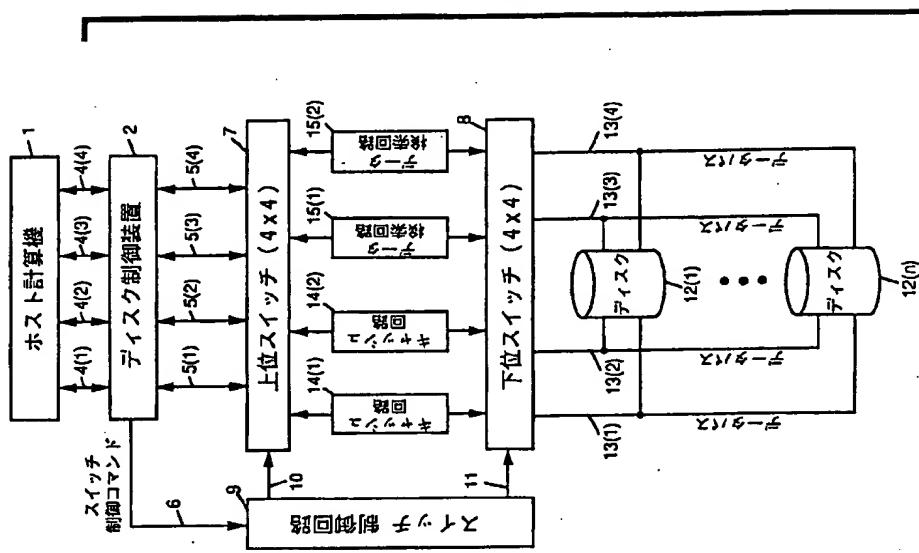


図2

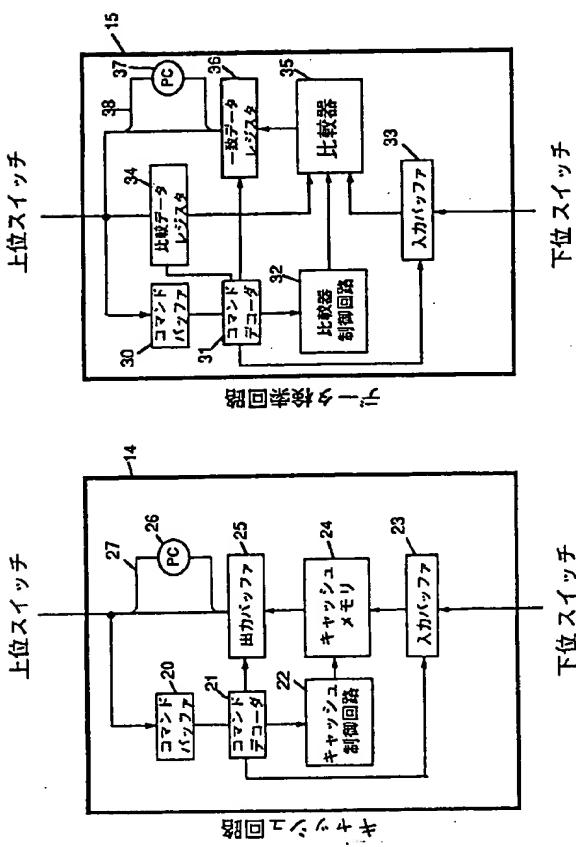


図3

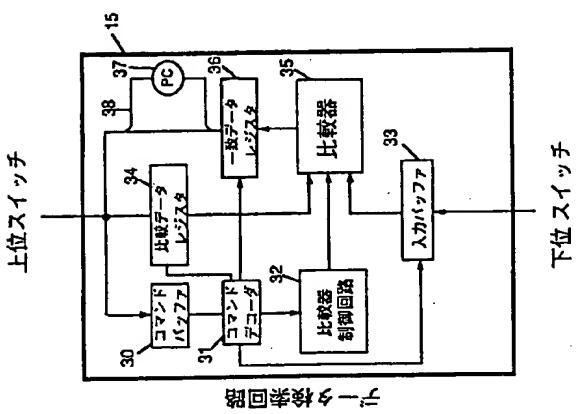


図10

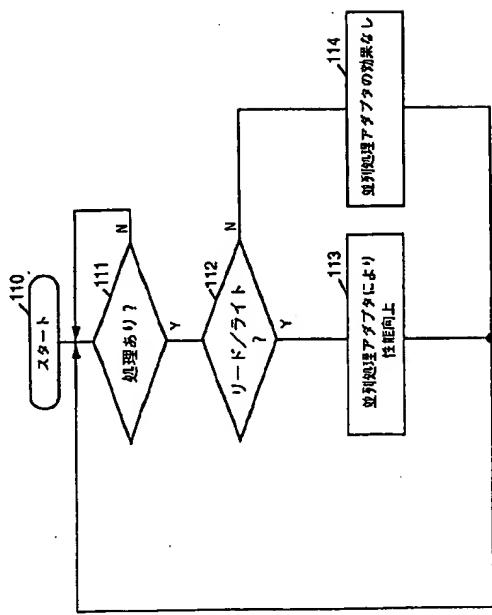
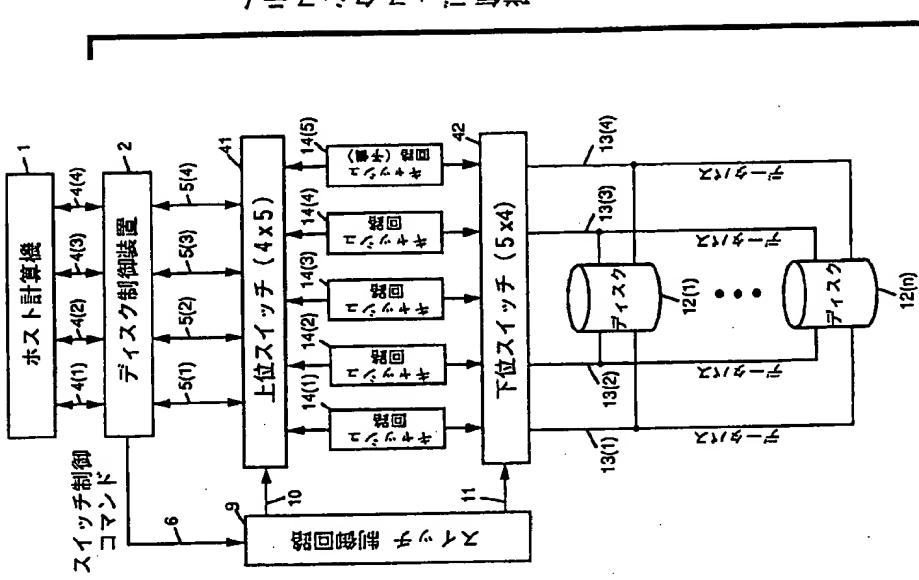
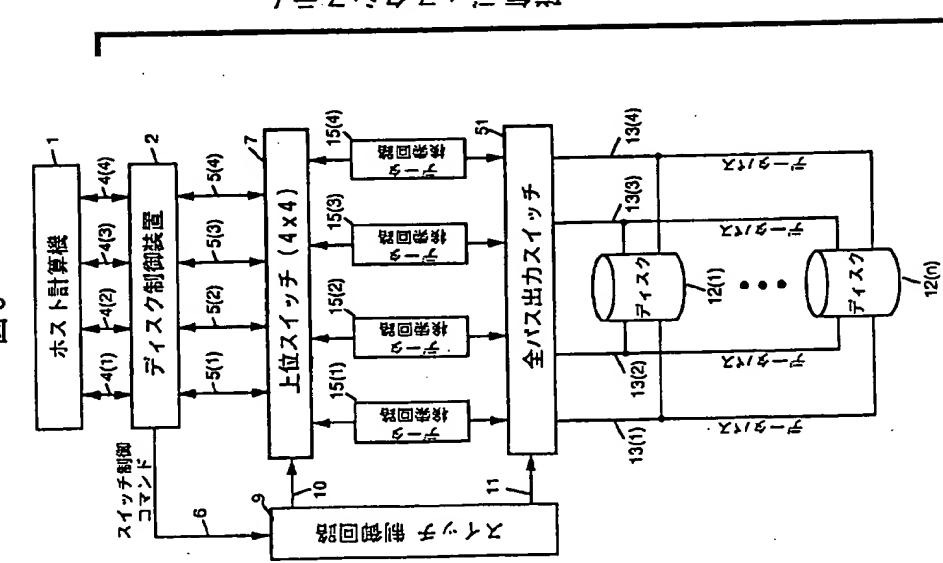


図4

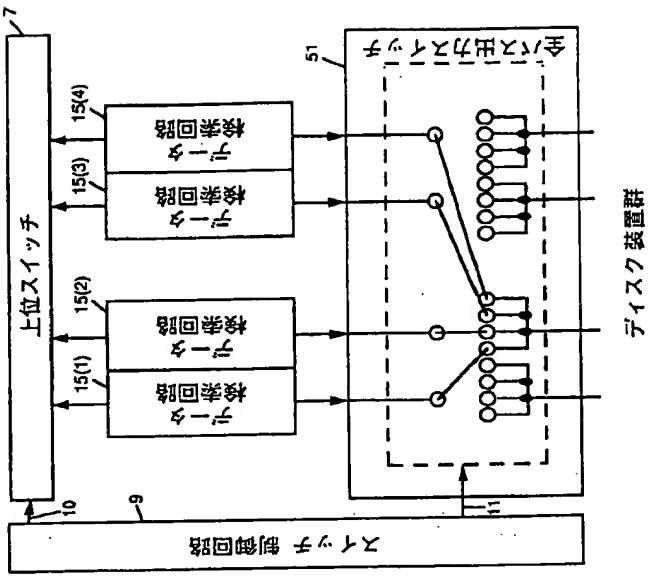


[図5]

[図6]

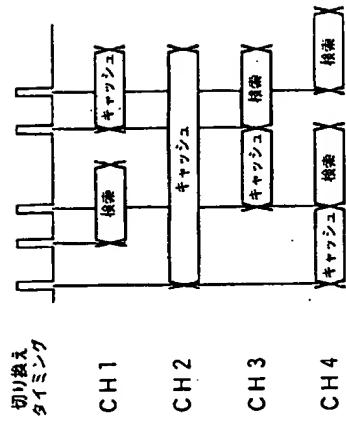


[図6]

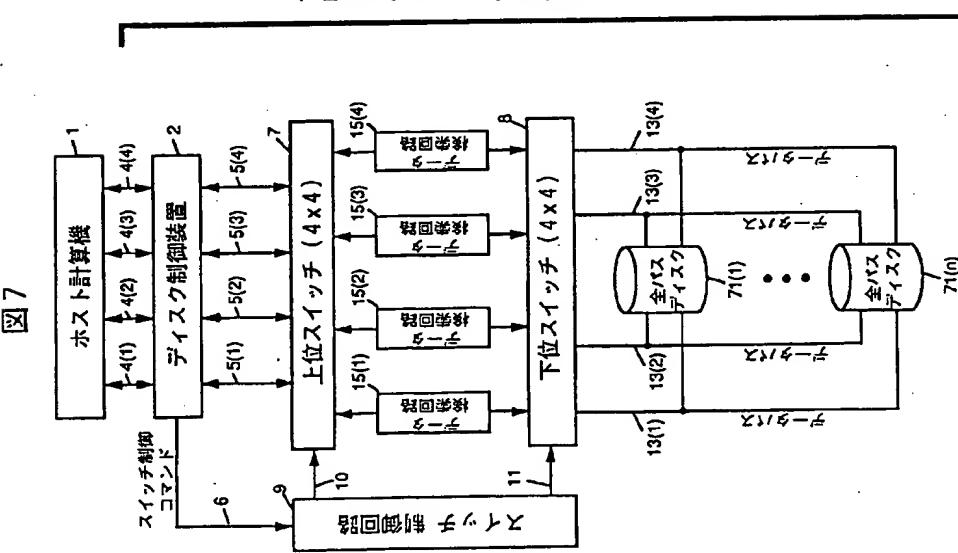


[図1.2]

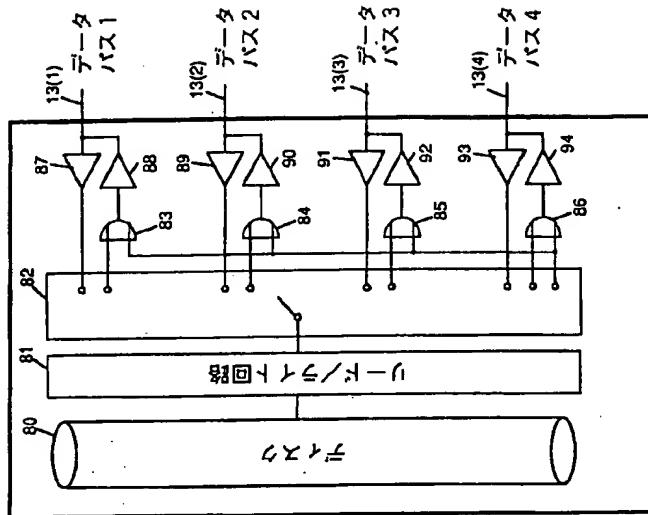
図1.2



四七

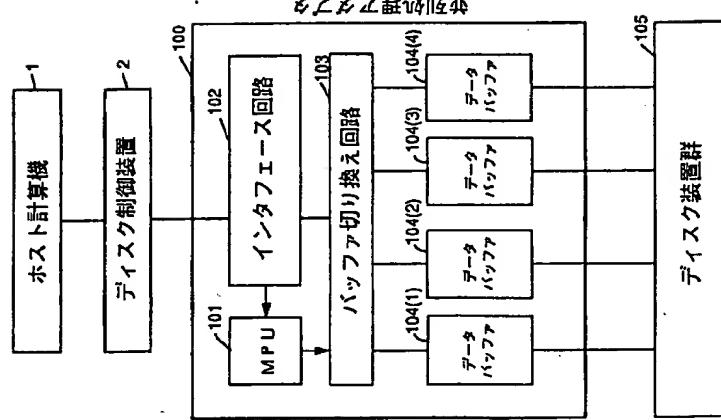


18



[図9]

図9



[図11]

図11

